Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Proceedings Series 2016/003

Pacific Region

Proceedings of the Pacific regional peer review on the Evaluation of the Internet Recreational Effort and Catch (iREC) Survey methods

June 2-3, 2015
Nanaimo, BC

Chairperson: Karen Hunter
Editor: Karen Hunter

Fisheries and Oceans Canada
Science Branch
3190 Hammond Bay Road
Nanaimo, BC V9T 6N7

February 2016

Canada

Foreword

The purpose of these Proceedings is to document the activities and key discussions of the meeting. The Proceedings may include research recommendations, uncertainties, and the rationale for decisions made during the meeting. Proceedings may also document when data, analyses or interpretations were reviewed and rejected on scientific grounds, including the reason(s) for rejection. As such, interpretations and opinions presented in this report individually may be factually incorrect or misleading, but are included to record as faithfully as possible what was considered at the meeting. No statements are to be taken as reflecting the conclusions of the meeting unless they are clearly identified as such. Moreover, further review may result in a change of conclusions where additional information was identified as relevant to the topics being considered, but not available in the timeframe of the meeting. In the rare case when there are formal dissenting views, these are also archived as Annexes to the Proceedings.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

© Her Majesty the Queen in Right of Canada, 2016
ISSN 1701-1280

Correct citation for this publication:

DFO. 2016. Proceedings of the Pacific regional peer review on the Evaluation of the Internet Recreational Effort and Catch (iREC) Survey methods; 2-3 June, 2015. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2016/003.

TABLE OF CONTENTS

SUMMARY iv

SOMMAIRE v

INTRODUCTION 1

WORKING PAPER #1 2

PRESENTATION OF WORKING PAPER 2

WRITTEN REVIEWS 2

Reviewer #1: Joe De Gisi 2

Reviewer #2: Carl Schwarz 3

GENERAL DISCUSSION 4

WORKING PAPER #2 6

PRESENTATION OF WORKING PAPER 6

WRITTEN REVIEWS 7

Reviewer #1: Joe De Gisi 7

Reviewer #2: Carl Schwarz 7

GENERAL DISCUSSION 8

APPENDIX A. TERMS OF REFERENCE 10

APPENDIX B. AGENDA 12

APPENDIX C. PARTICIPANTS 14

APPENDIX D. WRITTEN REVIEWS 15

SUMMARY

These Proceedings summarize the relevant discussions and key conclusions that resulted from the Fisheries and Oceans Canada (DFO), Canadian Science Advisory Secretariat (CSAS) Regional Peer Review meeting of June 2-3, 2015, at the Pacific Biological Station in Nanaimo, B.C. Two working papers were presented for peer review, focusing on the internet recreational effort and catch (iREC) survey and a method to calibrate iREC survey outputs to catch estimates generated from creel surveys.

In-person and web-based participation included Fisheries and Oceans Canada staff from the Science Sector and the Fisheries and Aquatic Management Sector, plus external participants from the commercial and recreational fishing sectors, academia, environmental non-governmental organizations, First Nations and the provincial government.

Current gaps in survey coverage used to monitor recreational fishing activity in Pacific region tidal waters (e.g. creel surveys) challenge Fisheries and Oceans Canada's ability to estimate total recreational catch and effort, and to conduct stock assessments or other analyses required by international agreements. The "internet recreational catch and effort" (iREC) survey was developed to provide catch and effort estimates for all combinations of area, month, fate, fishing method, and species harvested by the recreational sector. To minimize the effect of potential biases in iREC estimates, a calibration procedure was developed to relate iREC estimates and creel survey estimates in areas and times covered by a creel survey. Fisheries and Oceans Canada (DFO) Fisheries Management requested that DFO Science evaluate the iREC survey design and calibration methodology.

The conclusions and advice resulting from this review will be provided in the form of a Science Advisory Report supporting the iREC survey and calibration methodology for the generation of recreational fishing estimates for across area, time, species, fishing method and fate (kept and released).

The Science Advisory Report and supporting Research Document will be made publicly available on the Canadian Science Advisory Secretariat (CSAS) website.

Compte rendu de l'examen par les pairs régional du Pacifique sur l'évaluation des méthodes utilisées pour le sondage électronique de l'effort et des prises de la pêche récréative

SOMMAIRE

Le présent compte rendu résume les discussions et les principales conclusions de la réunion régionale d'examen par des pairs de Pêches et Océans Canada (MPO) et du Secrétariat canadien de consultation scientifique (SCCS) qui a eu lieu les 2 et 3 juin 2015 à la Station biologique du Pacifique de Nanaimo, en Colombie Britannique. Deux documents de travail ont été présentés aux fins d'examen par les pairs, portant sur le sondage électronique de l'effort et des prises de la pêche récréative (sondage iREC) et sur une méthode pour étalonner les résultats du sondage iREC en fonction des estimations de prises provenant des enquêtes par interrogation des pêcheurs.

Au nombre des participants qui ont assisté à la réunion en personne ou par conférence Web, on comptait des employés des secteurs des Sciences et de la Gestion des pêches et de l'aquaculture du MPO ainsi que des représentants secteurs de la pêche commerciale et récréative, des universités, d'organisations non gouvernementales de l'environnement, des Premières Nations et du gouvernement provincial.

Les lacunes actuelles dans la couverture des sondages utilisés pour surveiller les activités de pêche récréative dans les eaux de marée de la région du Pacifique (notamment des enquêtes par interrogation des pêcheurs) compromettent la capacité de Pêches et Océans Canada d'estimer les prises totales et l'effort de la pêche récréative, et d'effectuer les évaluations de stock ou les autres analyses exigées par les accords internationaux. Le sondage électronique de l'effort et des prises de la pêche récréative (iREC) a été élaboré dans le but de fournir une estimation des prises et de l'effort pour toutes les combinaisons de secteur, mois, devenir, méthode de pêche et espèce pêchée dans le cadre de la pêche récréative. Afin de limiter au maximum l'influence des biais potentiels sur les estimations de l'iREC, une procédure d'étalonnage a été mise en place pour mettre en relation les estimations de l'iREC et celles des estimations par interrogation des pêcheurs dans les secteurs et pour les périodes visés par une telle enquête. La direction de la Gestion des pêches de Pêches et Océans Canada a demandé au Secteur des sciences du MPO d'évaluer la conception de l'iREC et la méthode d'étalonnage.

Les conclusions et les recommandations découlant de cet examen seront présentées sous la forme d'un avis scientifique à l'appui du sondage iREC et de la méthode d'étalonnage pour générer des estimations de la production de la pêche récréative en croisant les secteurs, les périodes, les espèces, les méthodes de pêche et le devenir (conservé ou remis à l'eau).

L'avis scientifique et le document de recherche à l'appui seront rendus publics sur le site Web du Secrétariat canadien de consultation scientifique (SCCS).

INTRODUCTION

A Fisheries and Oceans Canada (DFO) Canadian Science Advisory Secretariat (CSAS), Regional Peer Review meeting was held on June 2-3, 2015 at the Pacific Biological Station in Nanaimo, BC, to review the internet recreational effort and catch (iREC) survey and a method to calibrate iREC survey outputs to catch estimates generated from creel surveys.

The Terms of Reference (TOR) for the science review (Appendix A) were developed in response to the Fisheries and Oceans Canada (DFO) Fisheries Management request to DFO Science to evaluate the iREC survey design and calibration methodology. Notifications of the science review and conditions for participation were sent to external representatives with relevant expertise from First Nations, commercial and recreational fishing sectors, environmental non-governmental organizations, provincial government and academia.

The following working papers (WP) were prepared and made available to meeting participants prior to the meeting:

1. Houtman, R., O'Brien, D.S., Komick, N., Rahme, A.H., and Hein, K. The "iREC" survey, an internet-based survey of recreational fishing effort and catch in British Columbia tidal waters: methodology and preliminary results. CSAP Working Paper 2014SFF01a.
2. O'Brien, D.S. and Houtman, R. Creel survey calibration of the iREC survey estimates: an integrated monitoring plan. CSAP Working Paper 2014SFF01b

The meeting Chair, Karen Hunter, welcomed participants and invited them to each introduce themselves, including those participating by Webinar. She reviewed the role of CSAS in the provision of peer-reviewed advice, and gave a general overview of the CSAS process. She discussed the role of participants, the purpose of the various resulting meeting publications (Science Advisory Report, Proceedings and Research Document), and the definition and process around achieving consensus decisions and advice. Everyone was invited to participate fully in the discussion and to contribute knowledge to the process, with the goal of delivering scientifically defensible conclusions and advice. It was confirmed with participants that all had received copies of the Terms of Reference, Agenda and the working papers.

The Chair reviewed the meeting's Terms of Reference (Appendix A) and Agenda (Appendix B). She then reviewed the ground rules and process for exchange, reminding participants that the meeting was a science review and not a consultation. The room was equipped with microphones to allow remote participation by web-based attendees, and in-person attendees were reminded to address comments and questions so they could be heard by those online.

Members were reminded that everyone at the meeting had equal standing as participants and that they were expected to contribute to the review process if they had information or questions relevant to the paper being discussed. In total, 35 people participated (listed in Appendix B). Nicholas Komick was identified as the Rapporteur.

Participants were informed that Carl Schwarz (Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC) and Joe De Gisi (BC Ministry of Forests, Lands and Resource Operations, Skeena Region) had been asked before the meeting to provide a detailed written review of the working papers to assist everyone attending. Participants were provided with copies of the written reviews prior to the meeting.

The conclusions and advice resulting from this review will be provided in the form of a Science Advisory Reports, which together with the resulting Research Document (the revised Working Paper) will be made publicly available on the Canadian Science Advisory Secretariat (CSAS) website.

WORKING PAPER #1

Working Paper: The "iREC" survey, an internet-based survey of recreational fishing effort and catch in British Columbia tidal waters: methodology and preliminary results by Robert Houtman, David O'Brien, Nicholas Komick, Ann Rahme, and Kristopher Hein (CSAP WP 2014SFFC1a)

Rapporteur: Nicholas Komick

Presenter(s): Robert Houtman, David O'Brien

PRESENTATION OF WORKING PAPER

Lead authors, Houtman and O'Brien, gave a presentation describing the internet recreational effort and catch survey. Limitations of traditional estimates of recreational catch in Pacific Region, based almost exclusively on creel surveys which only provide total effort and catch estimates for boat-based angling, and only cover certain areas and times, were outlined. The iREC survey is a new tool that uses an online survey, conducted monthly, to collect detailed recreational fishing catch and effort information from randomly selected Tidal Waters Sport Fishing License holders. The iREC survey was developed to fill creel survey coverage gaps and improve Fisheries and Oceans Canada's ability to conduct stock assessments or other analyses required by international agreements (e.g., Pacific Salmon Treaty [PST], International Pacific Halibut Commission [IPHC]) or for domestic management (sectoral allocation, Species at Risk Act [SARA], Fishery Monitoring and Catch Reporting [FMCR] Strategic Framework).

The authors explained that the survey had been operational as a pilot since July 2012, and was successful in producing catch and effort estimates for all areas, months, and fishing methods, for all species of interest encountered in the recreational sector. They explained the survey design, gave examples of the data outputs compared to creel, and suggested improvements that would improve survey design and possible reduce non-response bias. They recommended making licence purchaser aware of the legal requirement to participate in the survey should they be selected to the survey, at licence purchase. An explanation of various sources of bias that could affect a fisher-dependent survey such as iREC survey was provided. To address biases in iREC catch estimates a procedure to calibrate iREC estimates to creel-survey estimates has been developed and is presented in the second working paper of this review.

WRITTEN REVIEWS

Reviewer #1: Joe De Gisi

The reviewer congratulated the authors on completing this work and accomplishing the task of producing recreational fishing estimates for previously non-sampled areas and fisheries.

The reviewer provided several specific comments to improve the survey design including precision and uncertainty, participation, and the possibility of provincial collaboration in iREC.

The reviewer's full set of comments is available in Appendix D. Key comments included:

The reviewer made several requests of the authors to include additional information which he felt would better inform readers of the limitations and successes of iREC. There was limited reference to the extensive mail survey literature and the associated attributes of these kinds of surveys which the reviewer thought would be important to include. Information on the costs of managing the survey and cost-effectiveness of the method was requested.

The reviewer addressed the survey outputs by questioning what level of precision the survey is trying to achieve, and for what purpose. He requested that some explanation be provided to address how or if the survey meets management needs.

In response to comments in the working paper regarding collaboration between federal and provincial recreational fishery monitoring, the reviewer identified that the Provincial system for freshwater fishing licenses does not require email, and during online licence purchase, 30% to 40% of people don't provide email (it is optional). It was suggested that the discussion between provincial and federal partners should include capturing emails before the Province can consider doing an email-dependent survey.

Non-response to the survey was addressed by the reviewer. He recommended a telephone survey of non-responders be conducted to evaluate the magnitude of non-response bias and possible methods to improve response rates. He suggested the analysis include investigations into the rate of completion of the survey to understand non-response. For example, what occurs to partial responses, and are there populations of license holders that do not complete the survey? The reviewer believed that small additional investment on the quality of the survey interface could achieve a higher participation rate, perhaps by reducing dropout rate.

The reviewer indicated that a better description of the accuracy of the estimates provided in the working paper was required, specifically to clarify what the authors referred to as the 'reasonability of estimates'.

Bootstrapping techniques or other methods to understand characteristics of uncertainty of the estimates were recommended.

The reviewer discussed concerns regarding biases, and whether all bias should be handled by calibration. He suggested that major bias should be understood, without going down the 'rabbit hole' of never ending bias investigations. Because bias will likely change through time, if no effort is put into assessing bias, then the calibration may drift and be less accurate over time.

The reviewer commented that he thought un-calibrated iREC estimates were informative.

Author response

Authors thanked reviewers for providing helpful comments and a detailed technical review of the survey design.

They reminded the reviewers that information on creel survey uncertainty was included in the calibration methodology working paper which would be addressed later.

Authors accepted the suggestion to improve the survey interface, but noted that the interface was currently constrained by the design company.

Authors acknowledged that linking precision to management needs is important. They noted that there would be a workload cost with increasing sample size. However, there was some scope to more aggressively sample particular parts of the population. iREC results showed that there was similar precision to creel, which the Department has commonly used for management purposes.

Authors believe that biases that influence iREC will change over time. This guided them towards developing a calibration approach instead of trying to understand all bias. They agreed that key biases should be investigated. The review was helpful to devise a plan to address those.

The authors can include the comparative information related to costs of creel and iREC, as well as a section to address accuracy.

Reviewer #2: Carl Schwarz

Reviewer noted that the survey method had a lot of promise. The reviewer then discussed the most significant comments on the paper. More detailed comments are provided in the written review (Appendix E).

The review recommended that the paper needed to expand on the sampling approach to assure the reader that the survey works properly. Further, sample allocation is an important topic that was not covered well. The reviewer commented that the survey should be clear about precision goals and its distribution through the survey year. For example:

- What are the competing objectives for the survey and how do we allocate the sample?
- What is the desired precision rate on any particular estimate?
- Do you want consistent precision across months or do you prioritize particular estimates?

The reviewer identified that non-response bias is the major source of bias and is potentially a major issue for the survey. He recommended systematically investigating its importance. .

The reviewer explained that the approach used in the paper to estimate precision of estimates for combined Areas, summing variances of the Area-specific estimates, was incorrect due to the covariance of the response data for each of the individual areas. He indicated that the resulting error is likely small, but outlined an analytical procedure that would correctly estimate precision regardless of the scale of the estimate.

There was a strong recommendation from the reviewer to change the regression approach used to compare iREC to creel estimates to account for the uncertainty in both estimates.

The reviewer reminded the authors to be suspicious of the survey results and their accuracy. For example, when there is a low number of responses in a particular strata estimate, the estimate will be highly variable. He suggested using small area estimation methods to potentially achieve improved estimates.

Author Response

Authors thanked reviewer for providing a detailed technical review of the survey design. They agreed to address the majority of the reviewer's comments.

Regarding precision targeted at management objectives, the authors noted that optimal sample allocations would differ between species with different seasonal dynamics. They were not in a position to make specific decisions regarding precision that would meet particular management goals. The aim was to describe the survey design, its analysis and the results from the first 2.5 years of responses.

Authors noted that immediate selection of anglers at license purchase would help address the reviewer's concern about sample selection.

GENERAL DISCUSSION

The Chair invited questions or general comments from the participants on the presentations made by the authors and reviewers.

Participants provided commentary that suggested they were very pleased to see recreational fishing estimates for species, modes and areas never before sampled. There were a number of comments on technical aspects of the survey design. To organize these comments and discuss them systematically, the Chair referred to the objectives in the Terms of Reference and asked participants to share questions and comments linked to the working paper objectives. A list was generated which was then used to direct the discussion, and guided the decision on whether to accept or reject the paper.

Objective 1: *Document and evaluate the sampling design and analytical procedures, including the quantification of sampling error, used to estimate recreational catch from the iREC survey.*

Participants requested a description of the creel survey methodology and a table to summarize its sample rates. They recommended the paper include a section that compares and contrasts creel and iREC methods to understand what aspects are fishery dependent and fishery independent. Last, it was suggested to provide clarification that the iREC survey sampling is not stratified, the analysis is stratified, and to describe an example of the procedure.

Some comments about the application of the word bias were raised. Participants requested that authors clarify that bias applies to kept and released catch estimates, and that they avoid associating the word 'bias' (low or high) with either survey method (creel or iREC).

Participants agreed that it was important to highlight the differences between effort as estimated by iREC (licensee day) and most marine creel surveys (boat trips). There was some discussion that they were completely non-equivalent, and thus should not be compared, but the second working paper highlighted that a relationship between the two was useful for 'calibration' purposes.

There was a suggestion to include a brief literature review on mail and online survey response. DFO has completed mail surveys in the past which the authors could review and compare to the iREC response rate.

***Objective 2**: Quantitatively and qualitatively assess validity of the survey results across area, time, species, fishing method and fate (kept and released).*

Participants requested additional information on the survey allocation and sampling design. Descriptions of the basis for the sample allocation and strata (i.e. sample size across months, term and annual licences) used to generate survey estimates were requested. Participants suggested that the paper include discussion on how changes in survey design and allocation would impact precision, since different allocations could be used to provide improved precision on certain species. Participants made several suggestions on alternative stratifications (e.g. by age, home address components, whether or not the licence included a salmon stamp, licence purchase date, etc.) that could be used to correct for potential biases.

***Objective 3**: Identify and discuss potential sources of bias in these estimates.*

General comments on this objective included the need for more information on *what would* be an acceptable response rate given precision objectives (i.e. response rate).

The majority of the discussion focused on non-response oriented bias. Discussion included completing a literature review, incentives/disincentives for survey participants, incorporating a statement at license purchase specific to the legal requirement to complete the survey, translating the survey into multiple languages, improving the survey interface, and examining whether the time to complete the survey is a deterrent.

Additional suggestions were to evaluate existing data to determine the times and areas with the lowest response to encourage participation in those times and areas. Improved education about the survey was suggested as a method to improve response rates and reduce other biases. Another suggestion was to develop and implement diagnostics for identifying inappropriate response data.

Author response

The authors indicated they would add some paragraphs on survey methods to improve clarity around sampling approaches and rationales.

Authors responded that survey response rates were reasonable and results generally are consistent with known fishing activity (space and time) even with purposely no processing of the

response data to remove outliers. The authors agreed that diagnostics to remove obvious data entry errors or mistakes would be implemented.

Participants were reminded that the survey is mandatory. However, the survey takes a soft approach by sampling licencees a maximum of once a year (since April 2014). Data from 2012 demonstrated that survey fatigue was potentially problematic. Authors suggested maintaining minimal sampling of each licensee to avoid sampling fatigue and possible increases in non-response.

Incentives are not available as part of Department policy. They suggested that a possible cut in licence fees would be popular way to encourage participation. Authors noted that Washington State survey has low participation even with incentives.

The current allocation of the sample is not designed to meet particular precision targets beyond the rough estimation that given the response rates and variability within responses, the sample sizes chosen would provide a precision similar to that from creel surveys. However, a discussion about competing issues to identify sample sizes across months and between term and annual is needed. Precision of catch and effort estimates for a license type stratum depends not only on the number of selections of a particular license type, but on the response rate of selected people with that license type.

WORKING PAPER #2

Working Paper: Creel survey calibration of the IREC survey estimates: an integrated monitoring plan by David O'Brien and Robert Houtman. (CSAP Working Paper 2014SFF01b)

Rapporteur: Nicholas Komick

Presenter(s): David O'Brien

PRESENTATION OF WORKING PAPER

The lead author gave a presentation justifying calibration and outlining the statistical approach recommended to calibrate IREC estimates to creel survey estimates. The approach involves regressing IREC estimates against creel estimates for month-areas where the creel survey ran, and using parameters from those regressions for calibrating IREC estimates. The approach assumes that creel surveys produce less biased estimates on average than IREC estimates, so that calibration of IREC against the creel survey will provide less biased IREC estimates than un-calibrated IREC estimates. Currently, creel surveys conducted in tidal waters in the Pacific Region provide complete catch estimates only for boat-based angling. Thus, calibration of IREC estimates for methods other than boat-based angling would require either conduct of creel surveys designed to provide complete catch estimates for other methods, or application of boat-based calibration factors to the other methods.

Results of the calibration between IREC and creel boat-based angling estimates suggest that the calibration factor changes over time.

The data used in the regression would ideally represent a random selection of month-area combinations, to ensure that the regression parameters were not biased. However, such an approach is extremely impractical given the logistical and statistical challenges of conducting creel surveys in a more dispersed manner (reduced economy of scale), in less populous areas (staffing, travel costs) and areas and times with very low fishing effort (statistical challenges with sample size). Further, given the current approach to allocating creel survey coverage, to areas and times with relatively high effort and/or catches of Chinook Salmon and Halibut, the authors expected there to be a need to at least improve representation of the month-areas combinations

with low effort or catches. However, noting the number of creel estimates at relatively low values, the presenter argued that the distribution of effort and catches covered by the current creel allocation scheme was appropriate for calibration regressions.

WRITTEN REVIEWS

Reviewer #1: Joe De Giel

The reviewer commented that the purpose of the working paper was clear. He suggested there was a need for clarity when discussing bias and calibration.

The reviewer offered that the analysis that generated Figure 1 needed major revision because results were confounded between sample size and creel quality and that interpretation was essentially impossible. He suggested that if the analysis is to be presented, there needed to be more explanation about why the plot looks like it does.

The reviewer suggested that more work was required to represent uncertainty when comparing IREC and creel estimates. The reviewer noted that there are also likely sample size tradeoffs within the IREC: increasing the number of anglers surveyed per month may decrease the uncertainty of the estimates in the short term, but result in survey fatigue over time which could lower participation rates and re-inflate the uncertainty through time.

The reviewer noted that the Department's "Your Catch Counts" slogan for encouraging recreational fisheries catch monitoring may not be the best if you want to know if fishers didn't catch anything.

Author response

All comments provided are very helpful.

Reviewer #2: Carl Schwarz

Review of methodology

1. Use of ordinary least squares may be inappropriate

Creel survey estimates are used as the X variable in the calibration and there is considerable uncertainty in those estimates. The reviewer clarified that the least squares regression applied by the authors does not factor in error in the X values. He suggested a different regression method to handle uncertainty in both estimates.

The paper addresses how to estimate slope and intercept of comparison between IREC and creel. Least squares regression is not a correct approach because the independent variable values are estimates themselves, the "errors-in-variables" problem. The reviewer suggested that slopes and intercepts could be re-estimated annually using a Bayesian approach to combine prior years' results with the current year's dataset. Several years' data would be needed to assess whether there are real and persistent temporal, spatial, or species-specific differences between the calibration parameters. Additional research, data, and calibration runs are needed to improve the comparison and any resulting calibration.

2. Yearly/species calibration estimates

The reviewer noted that the correction factors change from year to year. He didn't believe the factor changes dramatically, and that the differences may be just variability. He suggested that a correction factor based on a longer time series, such as a running average of 3 years or other smoothing approaches, may provide the best correction factor.

Participants asked if there should be species specific calibration factors. The reviewer replied that statistically there is no problem to have a species specific correction factor, but he questioned whether this was practical. The reviewer re-iterated that the key issue is to tackle is non-response. For example, different responses by species as a result of recall bias or regulations could have an effect.

3. Designing creel survey to improve calibration

The reviewer suggested that authors deal with survey quality, especially in areas with limited creel coverage to improve the regression.

It will always be difficult to obtain good creel coverage where there is low creel effort. With low creel effort, there will be lower matching IREC responses for the specific anglers/with leading to significant expansion for the small number of responses. The reviewer suggested aggregating multiple areas to pull points away from zero. The reviewer noted the high cost associated with increasing creel sample sizes in areas with low effort.

The reviewer and participants agreed that it would be useful to incorporate information from other data sources to calibrate IREC estimates. Additional information would likely improve the process.

Author response

The authors acknowledged that all of the reviewers suggestions were extremely helpful, and they would incorporate the majority of the suggestions made.

GENERAL DISCUSSION

Similar to the first working paper, the Chair led participants through the objectives for the working paper, gathering comments and questions to be addressed by discussion. A list of issues was created and guided both the discussion and decision on whether to accept or reject the paper.

Objective 6: Develop methods to calibrate IREC estimates to the creel survey

Participants, in general, were somewhat skeptical about the calibration approach. Like the reviewers, participants struggled with the terminology used. They settled on calibration after comparing definitions of various terms. Participants rejected the least squares regression and accepted the revised method proposed by the reviewers. They suggested that the method should involve an iterative process to incorporate new data over time. Additionally, data sources that are regionally important, i.e. North Coast lodge logbook records, should be used to calibrate IREC in northern areas. Participants recommended further exploration of Bayesian approaches and further research into alternate calibration methods as additional years of data are gathered. Participants agreed to accept the approach on the understanding that revisions to the analysis would occur and would not change the overall outcome of the original calibration results presented at the meeting.

Considerable time was spent discussing the issue of low creel survey coverage (few samples or low catch) and its impact on estimates and calibration. Species with limited catch inherently have very sparse data. Participants asked if it was known where samples were needed to improve estimates. The suggestion to aggregate estimates for these species over space and time to improve catch estimates was deemed reasonable. Participants also suggested to use small area estimation to improve catch estimates in low data situations. As discussed by Di Schwarz, participants agreed that careful analysis of areas with low catch data was needed to handle high catch estimate variability among adjacent areas. They acknowledged that when aggregating areas, there is a tendency to apply the intercept multiple times and noted that

"sums of adjusted counts" (i.e. calibrated estimates) is not the same as "adjustment of the sum of the counts". The suggestion to force the regression through the origin could benefit the analysis by nullifying the error associated with summing values when aggregating areas.

There was a substantial amount of discussion about the need to recognize that the creel and other estimates against which the iREC survey would be calibrated are subject to biases themselves, rather than true 'standards' as implied by the term 'calibration'.

Participants raised the question of how or whether to calibrate catch estimates for areas and species without creel. The objectives of the working paper did not include this analysis. It was noted as an important research topic for future investigation.

Objective 5. *Explore and discuss options, including logistics and other potential constraints, to allocate creel survey or other methods to calibrate iREC estimates.*

Uncertainty in creel catch estimates was raised as the most important constraint to effective calibration of iREC estimates. Participants recommended additional data sources for calibrating iREC (e.g. North coast lodge books, guide log books). They reiterated that investigating non-response bias in iREC was absolutely needed, and increased certainty in iREC estimates could improve the calibration.

The authors provided an assessment of the quality of creel data used in the calibration applied to iREC estimates. Participants felt that the analysis required major revision. Participants requested to have information on the coverage of creel data used to calibrate iREC, and the degree to which creel data is fisher dependent (i.e. what proportion of creel interviews were fisher dependent). They were not clear on whether incomplete vs complete creel surveys were used and offered the suggestion to not use incomplete surveys, or to use a correction factor for incomplete creel months. A cut-off date or specific temporal and spatial scales was recommended when amassing creel data for calibration to avoid incomplete data sets.

A discussion ensued regarding what needed to be done to utilize iREC information most usefully. iREC provides data that could be used to support expanding traditional sampling methods for 'uncalibrated' species and fishing methods, but concern over feasibility and funding exist. There is a risk that creel survey methods and coverage might not be maintained due to funding constraints, and that any reduction in creel data would jeopardize the iREC calibration.

APPENDIX A. TERMS OF REFERENCE

Evaluation of the Internet Recreational Effort and Catch (iREC) Survey methods

Regional Peer Review Process – Pacific Region

June 2-3, 2015
Nanaimo, British Columbia

Chairperson: Karen Hunter

Context

Recreational catch estimates in the Pacific Region are based almost exclusively on creel surveys. These estimates only provide total effort and catch estimates for boat-based angling, and only cover certain areas and times. Such coverage gaps challenge Fisheries and Oceans Canada's ability to conduct stock assessments or other analyses required by international agreements (e.g., Pacific Salmon Treaty [PST], International Pacific Halibut Commission [IPHC]) or for domestic management (sectoral allocation, Species at Risk Act [SARA], Fishery Monitoring and Catch Reporting [FMCR] Strategic Framework).

The "internet recreational catch and effort" (iREC) survey, running as a pilot since July 2012, was developed to fill these coverage gaps in tidal waters, providing catch (and effort) estimates for all areas, months, and fishing methods, for all species of interest encountered by those methods. This survey, conducted monthly, involves using an internet survey to collect detailed catch and effort information from randomly selected license holders, and expanding those response data to represent the total catch and effort for all license holders.

To minimize the influence of potential biases in the responses on the final catch and effort estimates adopted as official catch, a calibration procedure is proposed in which the relationship between iREC estimates and creel survey estimates is used to calibrate iREC estimates in areas and times not covered by a creel survey.

Fisheries and Oceans Canada (DFO) Fisheries Management has requested that DFO Science evaluate the methods piloted to estimate recreational catch using the iREC survey, including the potential use of the creel survey to calibrate iREC estimates. In addition, if calibration is recommended, providing considerations for the allocation of creel surveys across months and areas has been requested. This assessment, and advice arising from this Canadian Science Advisory Secretariat (CSAS) Regional Peer Review (RPR), will be used to inform recreational catch monitoring planning, and as the basis for post-season catch estimates.

Objectives

The following two working papers will be reviewed and provide the basis for discussion and advice on the specific objectives outlined below:

Houtman, R., D. O'Brien, N. Komick. The "iREC" survey, an internet-based survey of recreational fishing effort and catch in British Columbia tidal waters: methodology and preliminary results. CSAP Working Paper 2014SFF01a

1. Document and evaluate the sampling design and analytical procedures, including the quantification of sampling error, used to estimate recreational catch from the iREC survey.
2. Quantitatively and qualitatively assess validity of the survey results across area, time, species, fishing method and fate (kept and released).
3. Identify and discuss potential sources of bias in these estimates.

O'Brien, D. and R. Houtman. Creel survey calibration of the iREC survey estimates: an integrated monitoring plan. CSAP Working Paper 2014SFF01b

4. Develop methods to calibrate iREC estimates to the creel survey.
5. Explore and discuss options, including logistics and other potential constraints, to allocate creel survey or other methods to calibrate iREC estimates.

Expected Publications

- Science Advisory Report (1)
- Proceedings
- Research Documents (2)

Participation

- Fisheries and Oceans Canada (DFO) Science & Management (shellfish, groundfish, pelagics and salmon)
- Academics
- Aboriginal communities/organizations
- Salmon and Halibut International Commissions
- British Columbia Sports Fishing Advisory Board
- Commercial Fishing Industry
- Environmental Non-government Organizations

APPENDIX B. AGENDA

Canadian Science Advisory Secretariat

Centre for Science Advice Pacific

Regional Peer Review Meeting

Evaluation of the Internet Recreational Effort and Catch (iREC) Survey methods

June 2 and 3, 2015

Pacific Biological Station
Nanaimo, BC

Chair: Karen Hunter

DAY 1 - Tuesday, June 2nd 2015

Time	Subject	Presenter
0900	Introductions Review Agenda & Housekeeping CSAS Overview and Procedures Review Terms of Reference	Chair
0930	Presentation of Working Paper #1	Authors: Houtman, O'Brien, Komick, Rahme, Hein
1000	Presentation of Written Reviews – Working Paper #1	Reviewers: DeGisi and Schwarz + Authors
1030	***Break***	
1050	Identification of Key Issues for Group Discussion	Participants
1130	Discussion & Resolution of Technical Issues	Participants
12:00	***Lunch Break***	
1300	Discussion & Resolution of Results & Conclusions	Participants
1430	Develop Consensus on Paper Acceptability & Agreed-upon Revisions	Participants
1445	***Break***	
1500	Presentation of Working Paper #2	Authors: O'Brien, Houtman,
1530	Presentation of Written Reviews – Working Paper #2	Reviewers: DeGisi and Schwarz + Authors
1600	Identification of Key Issues for Group Discussion, Paper #2	Participants
1700	Adjourn for the Day	

DAY 2 - Wednesday, June 3rd 2015

Time	Subject	Presenter
0900	Introductions Review Agenda & Housekeeping Review Status of Day 1	Chair
915	Discussion & Resolution of Results & Conclusions, Paper #2	Participants
1000	Develop Consensus on Paper Acceptability & Agreed-upon Revisions, Paper #2	Participants
1030	***Break***	
1100	Science Advisory Report (SAR) Develop consensus on the following for inclusion: • Sources of Uncertainty • Results & Conclusions Additional advice to Management (as warranted)	Participants
1200	***Lunch Break***	
1300	Science Advisory Report (SAR) • Continued	Participants
1445	***Break***	
1500	Next Steps – Chair to review • SAR review/approval process and timelines • Research Document & Proceedings timelines • Other follow-up or commitments (*as necessary)*	Chair
1545	Other Business arising from the review	Chair & Participants
1600	***Adjourn meeting***	

APPENDIX C. PARTICIPANTS

Last Name	First Name	Affiliation
Adams	Devona	DFO Fisheries Management
Beaith	Brad	DFO Fisheries Management
Campbell	Kelsey	A'Tlegay Fisheries Society
DeGisi	Joe	Province of British Columbia
Eros	Carole	DFO Fisheries Management
Fong	Ken	DFO Science
Fraser	Kathy	DFO Science
Gale	Rupert	Sports Fish Advisory Board
Hill	Aaron	Marine Conservation Caucus
Houtman	Rob	DFO Science
Hunter	Karen	DFO Science
Komick	Nick	DFO Science
Laliberte	Bernette	Cowichan Tribes
Lane	Jim	Uu-a-thluk/Nuu-chah-nulth Tribal Council
Luedke	Wilf	DFO Science
MacDougall	Lesley	DFO Science
Maher	Jordan	Q'ul-lhanumutsun Aquatic Resources Society
Nicklin	Pete	Fraser River Aboriginal Secretariat
O'Brien	Dave	DFO Science
Parken	Chuck	DFO Science
Rahme	Ann	DFO Science
Rutherford	Kate	DFO Science
Rutherford	Dennis	DFO Science
Sawada	Joel	DFO Science
Schwarz	Carl	Simon Fraser University
Scroggie	Jamie	DFO Fisheries Management
Staley	Mike	Fraser River Aboriginal Secretariat
Tadey	Rob	DFO Fisheries Management
Tadey	Joe	DFO Science
Tompkins	Arlene	DFO Science
Webster	Ray	International Pacific Halibut Commission
Winther	Ivan	DFO Science

APPENDIX D. WRITTEN REVIEWS

Reviewer: Joe De Gisi

BC Ministry of Forests, Lands and Resource Operations
Skeena Region
3726 Alfred Ave
Smithers, BC

CSAS Working Papers: 2014SFF01a and 2014SFF01b

Is the purpose of the working paper clearly stated?

As with the companion paper, this document provides a response to a request from the Fisheries Management Branch for advice about the use for management purposes of the iREC survey. The iREC is an internet-based survey which is intended to estimate monthly effort and catch in tidal recreational fisheries of the Pacific Region, for all species, fishing methods, and management areas.

Stock assessment and management activities presently require these types of information in the context of allocation, exploitation rate estimation, and catch reporting requirements. This paper states that it is an attempt to address two specific components of the request: the development of basic methods to calibrate iREC, and how creel resources could be allocated to most effectively achieve the calibration.

Are the data and methods explained in sufficient detail to properly evaluate the conclusions? Are the data and methods adequate to support the conclusions?

The conclusions, couched as recommendations to Management, are mainly provided in Section 4.1 (p. 17). These are as follows.

(A) Creel survey times/places for iREC calibration should continue to be prioritized based on recreational fishery effort, and catch of key species.

This recommendation is based upon achieving multiple priorities of

1. maintaining recent traditional estimates of effort and catch for the dominant actively-managed fisheries,
2. acquiring the most valuable data for calibration, and
3. collecting other bio-information needed for specific management purposes.

Although the report states (p. 8) that the iREC has been used to help prioritize the allocation of resources, it appears that the pre-existing focus on area-month combinations with greatest recreational effort, and highest retained catch of the most-sought species meets all three priorities and should be continued. This appears to be a warranted conclusion from the results presented.

(B) Standardizing sample rates and analytical approaches, and ensuring complete spatial/temporal coverage within the identified month/area/method combinations, will improve calibration value.

This recommendation is intuitive but not strongly justified by the analyses presented in the document. Figure 1 shows how the slopes of the calibration-type regressions stabilize as data pairs are added, in order of decreasing "quality" (1a) and increasing "quality" (1b). For the most part, the slopes stabilize only slightly earlier in plot 1a, suggesting (as the report does) that for calibration with this type of data the number of estimate pairs is more important than ranked estimate quality. This may be due to the way that quality was ranked, possibly not capturing

quality in the sense of accuracy or precision, or it may simply be high sampling error in either type of estimate. Summarizing the relative precision of each rank group (1 to 7) of traditional estimate could be informative. Regardless, these rankings of quality involve the factors that can be controlled in traditional recreational fishery surveys. Assuming that resources for traditional surveys are limited, there is a tradeoff between the quality and the quantity of surveys that can be made, and more analysis is needed to justify the recommendation provided above. It appears prudent to follow this recommendation to collect additional year's data and re-evaluate.

(C) Where regression parameter estimates differ significantly from intercept of zero and slope of 1, traditional survey estimates should be used to calibrate iREC.

It is clear that as a general rule for iREC, bias correction or calibration is required, and that the primary questions are how calibration slopes and intercepts should be estimated. First, the use of simple least squares regression is not appropriate, because the data display the errors-in-variables problem. Slopes and intercepts can be re-estimated annually using a Bayesian approach to combine prior years' results with the current year's dataset. Several years' data will be needed to assess whether there are real and persistent temporal, spatial, or species-specific differences between the calibration parameters.

(D) Annual catch estimates should be made as the sum of traditional estimates in the month/areas where they are performed and calibrated iREC estimates for other times/areas.

If possible, the document should provide an explanation of how uncertainty will be estimated using this approach. Uncertainty will continue to be high (relative to the estimate) for area/month combinations where activity and catch are lower but these may contribute relatively little to the overall uncertainty.

If the document presents advice to decision-makers, are the recommendations provided in a useable form, and does the advice reflect the uncertainty in the data, analysis or process?

The paper does provide useable advice and recommendations which take explicit account of uncertainty. Management presumably has objectives for relative precision, and there are tradeoffs involved in allocating budget resources to reducing the uncertainty. The iREC is inexpensive for the quantity of information gained, and generates a large number of area/month/species estimates which would be unobtainable otherwise with available resources. Calibration by comparison to traditional recreational fishery surveys may reduce the bias (on average) of the iREC estimates but potentially reduce the precision due to the uncertainty in the calibration regression parameters. There are also likely sample size tradeoffs within the iREC; increasing the number of anglers surveyed per month may decreases the uncertainty of the estimates in the short term but result in survey fatigue over time which could lower participation rates and re-inflate the uncertainty through time. It is difficult for this paper to provide definitive advice given the short time over which the iREC has been in existence; more data will allow better analyses

Can you suggest additional areas of research that are needed to improve our assessment abilities?

Two areas of research are suggested. First, more years of data and analysis will be needed to determine how the calibrations should be estimated, in particular which species (if any) share common calibration relationships and how these may vary in space and time. Second, a better understanding of the sources of upward bias in the iREC estimates is desirable. Although it may not be possible to eliminate the sources of bias, not knowing the its primary cause(s) is risky. Changes in the causes, possibly leading calibrations to drift, cannot be monitored if they are unknown.

Date: 2015-06-03

Reviewer: Carl James Schwarz

Department of Statistics and Actuarial Science
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6

CSAS Working Papers: 2014SFF01a and 2014SFF01b

1. Introduction

This is a review of two working papers for the Canadian Science Advisory Secretariat (CSAS) on the internet based survey of recreational fishing effort and catch. Both documents were provided by email in the mid- to end-of May 2015.

The purpose of the *iRec Survey* is to provide estimates of recreational effort, catch, and harvest for in BC tidal waters.

All licenses are now sold via an electronic system which greatly simplifies the acquision of basic license data. With the near ubiquity of internet availability for the population, it is hope that an electronic capture of the data will provide a cost effective method to obtain high quality estimates.

2. Review of 2014SFF01a – iRec design and analysis

2.1 Review of current survey design.

While an annual license is valid for the entire fiscal year, they can be purchased at any time. Consequently, for any survey month, the pool of license holders is first divided into those licenses that were purchased prior to the survey month (Annual-Pre), and those licenses that were purchased during the survey month (Annual-Post). A license that was in the Annual-Post group "transfers" to the Annual-Pre group at the start of the next month.

Both groups of licenses are sampled at the same rate (p. 17), with a summary of the sample allocation in Table 3. It would have been helpful in Table 3 to show the number of licenses in the frame prior to sampling. Note that according to Table 3, it does not appear possible to purchase a license prior to the start of the fiscal year for which it is valid, i.e. no Annual-Pre licenses are available for April of any year.

Term licenses that are valid for at least part of the survey month are sampled at 25%. It was unclear from the report if a term license whose term spans a month boundary are part of the frame for both months.

Currently, a license which was selected in any month is excluded from being sampled in subsequent months (until the end of the fiscal year, when all licenses expire). The report (page 18) indicates that an adjustment is made to avoid potential bias caused by this non-replacement sampling

> "To ensure Annual-Pre samples are representative with respect to license purchase date, licenses available for the Annual-Pre sample selection are stratified by month and sampled proportionate to the fraction of total annual licenses sold in that month regardless of earlier selections" (p.18).

I'm not clear exactly what is done here – some additional explanation will be required.

Some adjustment is made for the same person holding more than one license during a survey month, but these are minor.

Holders of licenses selected in the survey month are contacted by email and asked to complete the survey online. Up to two reminder emails can be send to license holders who have not responded (p. 20).

The individualized emails link to the data collection pages where days fished, catch, kept, released, method of fishing, etc. are collected. Because juveniles are not surveyed separately, license holders also indicate if juveniles were fishing with them during the survey period. It is not clear to me that records of catch for shellfish will be reliable – it is hard to believe that a license holder would remember how many clams were harvested.

The analysis uses simple expansion estimators, i.e. the mean response value is inflated by the population size. Separate estimates are obtained for combinations of license type and term (p. 25). Some of these variable are known when the license is purchased and are static (e.g. residence and age), but some of these variables can change (Annual-Post to Annual-Pre). The report calls these divisions "strata" (p.25), but the term "stratum" is usually reserved in survey-sampling for a fixed-unchanging aspect of a population.

The estimated variance for the mean of the response (p. 27) is based on a simple-random sample in a stratified design. Similarly, the estimated variance for the overall mean (and total) also is that from a simple-random sample in a stratified design. A large sample confidence interval is used (but the df for the *t*-distribution was never specified).

The report does not specify how estimates at larger scales are obtained. For example, it would not be appropriate to simply add the estimates from each area because some license holders may have fished in multiple areas. Rather, the sum of the catch over the larger scale must be determined for each respondent and this derived variable used in the estimation formulae.

Computations are done using *R* – it is unclear if this is custom code, or are standard survey packages (such as *survey*) being used.

2.2 Theoretical justification of proposed methodology is needed.

The report did not make it clear why they implemented their sampling design in the way they did. Is there some limitation of the licensing system that makes it difficult to sample licenses directly (as noted below)?

The current design does not fit one of the standard survey designs traditionally seen in textbooks, so some theoretical investigation is needed to ensure that its properties are known.

I was unable, for example, to figure out the probability of selection of a license-month. For example, the first two equations on page 26 implicitly assume that all license-months in

category have the same probability of selection ( ), but licenses that were previously selected have a probability of 0 of being selected in this month, and licenses purchased earlier in the year would appear to have a lower chance of being selected than licenses purchased later in the year (except for the note on page 18 cited earlier which was not very clear).

Similarly, the estimation formula make an implicit assumption that all probability of selection are independent within and across categories and months – it is unclear that this is true.

Without this basic information, it is unclear if the estimation methods used in the report are appropriate. Given the magnitude of the licence sales and size of the samples, I suspect that any biases from unequal probabilities of selection and estimation methods are small, but would like some assurances. Non-response bias is likely the largest source of bias and likely swamps all other sources of bias. The report needs to step back and do some theoretical investigation of their design.

It may be beneficial and conceptually much easier to visualize the process if the survey design was viewed differently, namely as two-stage design with multiple levels of stratification. The ultimate survey unit is the licence-month, i.e. a combination of licence x (active) month of the year. These survey-units are, by definitions, found in clusters, namely the licence. Once a cluster (licence) is selected, the survey period can be selected from the active months for the licence. By looking at the survey in this manner, all of the problems associated with ensuring that the same licence is not selected for more than one survey months are automatically avoided, a great deal more flexibility about sample allocation is possible compared to the previous approach, and the theoretical properties of the design are now easily studied.

In the first stage of the design, the population is divided into (at least) 8 classes (Annual/Term x resident/non-resident x Adult/Senior). A second classification corresponding to the month purchased (12 classes corresponding April, May, June ..., March). There are then 8 x 13 strata where each licence falls into one and only one stratum (known in advance). A final stratum for the few paper licences is easily accommodated. The number of licences in each stratum is known.

A sample of licenses is selected from each of the 8 x 13 strata using some probabilistic scheme. Here the probability of selection can be easily varied by strata to accommodate management objectives and precision goals. Licence holders do not have to be sampled with equal probability from each stratum (e.g. some sort of unequal probability sampling could be used within stratum depending on other attributes of the licence). For example, unlike in the current design, residents and non-residents can be sampled with different probabilities for specified objectives. At the moment, no use appears to be made of the salmon stamp – this could be used to select licences with a different probability to obtain estimates of the salmon harvest with a higher precision. The location of purchase of licence is presumably available – it may be that this can used to modify the probability of selection to ensure that certain areas have adequate sample sizes.

Once a license has been selected, then the survey period(s) for each licence (month(s) for which reporting is required) is selected – refer to Figure 1. Currently, each licence can only be sampled in one month, but there are no theoretical reasons why multiple months cannot be selected. The survey period (month) can again be selected with unequal probability to ensure, for example, that month of purchase are selected with higher probability, or that months in peak effort times, are also selected with higher probability. This is subsampling with (as currently done) only one survey period selected for each licence.

There are many response variables – the catch and effort across the management areas, fishing methods, species (and combinations thereof). Once the estimation methods are established for one response variable, then similar methods can be used for other similar response variables or aggregates (e.g. total catch over all management areas in that month) with out much difficulty.

The probability of selecting any licence-month is now easily computed and the standard methods for two-stage sampling can be employed (e.g. Thompson, 2012). As well, standard software for survey analysis (e.g. the survey package from *R*) can now be used rather than writing bespoke code.

For example, consider estimating the total catch in a particular month in one stratum. This is an example of domain estimation with the domain being the month of interest for a licence holder. Not all licences from a stratum were sub-sampled for that month, but the number of active licences for that month in the stratum are known. We can now treat the month as a sub-population where the population size of the sub-population is known, and the standard

estimators and variance formulas can be used (see, for example, Thompson, Section 3.6 for the case of a SRS, the extension to an unequal probability design is straight forward).

Some care is needed when aggregates are estimated (e.g. total catch over all management areas). Estimates should not be simply summed over the individual management areas to get the estimate for the larger aggregate, because the same survey units are repeatedly used in each of the smaller areas. Rather an aggregate response should be computed for each respondent and then the standard survey methods used on the aggregate response.

Because each license is now only selected for one month, the estimates of the monthly totals are independent and the yearly totals/se are easily found by adding the individual monthly totals. If licences can be surveyed in multiple months, then a more complex method must be used to get yearly totals to account for the multiple measurements on the same licence. As with all sub-sampling designs, the first step is to work backward to estimate the cluster (licence) total and then use the estimated totals in estimating stratum means and totals. For example, a simple way is to expand the reported catch in the sampled months to the yearly total by 365/number of fishing days available recorded for the licence. This is now an estimate of the total catch in that domain for the year for that licence.

Once estimates are found for each stratum, the estimated total/se combined over strata are easily found because no licence can be found in multiple strata and so all strata estimates are independent of each other.

2.3 Use standard survey terminology.

The terminology used in this report is a mixture of standard survey terms and common usage terms. This can be confusing to a reader accustomed to surveys. For example, p. 16, "The participant selection consists of three sub-samples". The sub-samples here refer to the Annual-Pre, the Annual-Post, and Term licence holders. The Term licences are a stratum (term licences cannot change into annual licences), but licences switch from Annual-Post to Annual-Pre and so these are not strata, nor are they sub-samples. The term *sub-samples* is traditionally is used in survey-sampling for two-stage sampling design. Later in the document (page 25), the report indicates "The calculations are stratified by licence type ... and subsample (Annual-Pre, Annual-Post, and Term)" which is very confusing.

There are multiple response variables (total catch, kept, released, effort) for each management area and fishing method etc. Rather than indexing these multiple responses using subscripts, just create a generic response variable Y because the estimation formula are identical for the different responses and this will reduce clutter in the report and make it easier to read. For example, the estimating equations for effort are identical to those for catch and do not need to be repeated. Similarly, aggregate responses (e.g. total catch over all management areas) should be created first for each respondent and then the aggregate used with the standard survey formulae.

The report's notation (e.g. p. 26) doesn't differential between estimators and population values very well. For example, the reader needs to carefully distinguish between the estimated catch and the population value thereof, the estimated variance of catch and the population value thereof. Standard notation (e.g. circumflexes) should be used to indicate which quantities are estimates and which are population value.

2.4 Sample Allocation

I could find no rationale in the document for the sample allocation (Table 3). The Annual-Pre licence sample size is approximately uniform across the year, but the Annual-Post and Term

licence sample appear to more heavily concentrated in the months with higher effort. Why is a similar allocation used for the Annual Pre licences?

There are number of competing objectives for this survey, not all of which have equal priority. For example, are estimates of salmon harvest a higher priority than estimates of shellfish harvest? Are estimates needed with equal precision for each month (favouring a more uniform allocation of sampling effort) or is the total yearly harvest of primary interest (favouring allocations to months with more effort). Some discussion is needed of the overall objectives to inform the choice of sample allocation.

With the very small marginal cost of an email and a fully electronic data collection, why isn't every licence holder selected at some point over the year rather than trying to only select certain holders? Then the sample allocation problem reduces considerably to deciding in just which month each licence holder should be sampled.

2.5 Non-response bias

The largest problem facing Creel is potential non-response bias given that only about 30% of selected respondents actually respond to the survey (Table 3). The report identifies some reasons for non-response (p. 23) including non-delivery due to defective email addresses, and spam filtering. If these occur at random, then no biases are introduced to the estimates – however, it is impossible to ascertain if this occurred because there is virtually no information available on spam filtering. A much more serious concern is non-response and response bias by licence holders for various reasons. For example, avid fishers may be over-represented in the responses which will lead to a positive bias in estimates; some fishers may try and "game" the system by not responding or underreporting catch; recall bias may give catch numbers that can be biased in either direction; etc (p. 66). Respondents may be reluctant to report catches during closures or above limits given that a licence number is associated with responses.

The report does not investigate the potential response bias by comparing demographic characteristics (e.g. age) and survey responses from early responders to responses from later responders (e.g. after one or more reminders). For example, Groves et al. (2009) illustrates how to estimate potential biases by comparing responses from successive waves of respondents. A separate survey of non-responders (e.g. by telephone) may be another way to try and identify reasons for non-response and potential biases. Additional effort is needed in this area to ensure that non-response bias is not severe.

Is the birthdate captured during the licence purchase? If so, then you can use the current respondents to see if the distribution of age is comparable to those of non-respondents. Perhaps older licence holders are more reluctant to respond?

A common way to (partially) deal with response bias is to reweight the data to account for known demographic biases. For example, older respondents may not be as comfortable in responding electronically and so may have a high non-response rate than younger respondents. Consequently, reweighting the responses to account for known demographic imbalance can be used.

Incentives to respond could also be used to increase the response rate. For example, if all licence holders were surveyed (but in only one month), a prize draw from all licence holders that respond may encourage holders to respond.

The computerized selection and data collection makes it straightforward to run "experiments" to investigate many aspect of the survey. For example, in one year,

every second selected licence holders would use the current system. The other holders could receive a simplified reporting system such as a simple list of "Did you fish in the following set of

require during the survey period regardless if catching any fish" (this would give an estimate of the participation rate in a particular area-month combination) and then just request the total number of fish caught, kept, released during the month rather than filling out the daily catch record. This experiment would investigate if the respondents found the survey too onerous to respond. [By the way, presumably statistics are available on the number of surveys started but not finished?] Similarly, an experiment could be run to compare the internet survey vs. a texting-based survey. A prioritized list of questions should be prepared for further investigation.

6.6 Consider small area estimation.

The survey attempts to find estimates a very fine scale and each estimate is derived independently for each response. Consequently, many estimates are based on very small number of respondents and highly variable. Even worse, the estimated standard errors are also high variable. For example, small changes in the data may change the estimates from 2 to a large number and change the standard error in similar ways. To obtain better estimates for small scales, a number of techniques have been developed that 'borrow strength' from surrounding areas. One approach, referred to as synthetic estimation, makes use of estimates and auxiliary information from the larger area to adjust estimates for the small area of interest. Composite estimators use a weighted average of the synthetic and direct estimators. A variety of model-based methods using auxiliary information from the small area and larger population have also been introduced. Ghosh and Rao (1994) provide a review and appraisal of small-area estimation methods. This may be an area for further research down the road.

6.7 Comparison with other surveys.

The paper compares the results from the *Rec* survey to those from other surveys (primarily creel surveys). A direct comparison is not possible because the creel surveys are often time limited (only in months with large effort) and in specific areas (only in areas with large effort).

Results are presented mostly using the *heat-maps* which I found effective to view a large amount of information at a high level. As noted in the report, these heat-maps were consistent with expectation indicating that there are unlikely to be any large scale relative problems in the estimates, but these maps cannot be used to investigate biases in absolute numbers.

Table 6 presents comparison of catch rates from *Rec* and those from creel surveys. These estimates appear to be comparable (except for shellfish) but no standard errors of either estimates were reported which makes it difficult to interpret any differences.

Table 5 presents selected estimates of catch to be compared to creel estimates. The report does not indicate how the *Rec* estimates were obtained. As noted earlier, the separate estimates by management area must be combined carefully because the same respondents are used multiple times. For example, consider the first line for Chinook Salmon in Table 5. The estimated catch in 2018 on the north coast is 327/88 (se 3809) and the estimated catch in 2018 on the south coast is 3007/59 (se 5855). The estimated sum appears to be found as

327/sqrt(3809^2 + 5855^2) = 39e-6/7 and the estimated se appears to be found as . This needs to be examined as the same respondents would be used in both estimates and so the estimates are not independent. [It may turn out that virtually no respondents fished in both coasts and so these observed results may be an artifact of the dataset.] The appropriate aggregate must be obtained for each respondent, i.e. create a combined north and south coast for each respondent, and then use the aggregate in the estimating equations.

Simple sums over months is currently ok because each respondent is selected for only once, but the standard errors when summing over months need more thought because of the way that

respondents were selected only once which induces some funky correlation among the monthly estimates.

Estimates from the *iRec* tend to be larger on the South Coast compared to the estimates from the creel survey. The authors attribute this to the difference in survey coverage between the two survey methods. It would be helpful to get some (rather admitted coarse) estimates of the relative coverage of the creel survey vs. the *iRec* survey to see if this matches the degree of difference in the estimates.

The authors used a regression approach to investigate the relationship between the estimates of the two surveys. I would have liked to see some plots of the estimates with error bars in both the vertical and horizontal directions (both surveys are only estimates). I don't know how much I would true the estimates slopes and p-value because both the *X*-variable (creel survey estimates) and *Y*-variable (*iRec* estimates) have considerable uncertainty. This error-in-variables approach is known to typically attenuate slopes (i.e. bring closer to 0) and the ordinary p-values from a simple regression are not reliable. I don't have a feeling on how bad this could be because it is not clear from the tables the size of the relative error in the estimates. The very low value of R^2 for many relationships indicates tremendous scatter in the estimates.

2.8. Investigate other survey modes.

While this survey uses the internet to gather data from respondents, an alternate approach would be to use text messaging and smart phones rather than email and desktop/laptop browsers. Rather than sending an email which is easily ignored, the licence holder would receive a text message and could report their catch/ harvest, etc. also via a text message or a simple smartphone application.

Or, an application could be developed that allows purchases of licences and recording catch on the fly.

These are examples as noted previously of the types of experiments possible.

3. Review of 2014SFF01b – Calibration

3.1 Review of methodology

Because the data from the survey is from voluntary responses from licence holders, it is quite conceivable that non-response bias could be extensive. Consequently, the results from the *iRec* survey for selected area-month-species combinations and species are compared to a respondent-independent creel survey.

While the term "calibration" is used here, this has a specific mean in survey-sampling where the survey weights are adjusted to force estimates of auxiliary variable to match known auxiliary variable totals (Samdal, 2007). The informal use of calibration will be used in the remainder of the document.

The methodology used here is to plot the estimates from *iRec* against those from the creel-survey and then fit an ordinary least-squares line to the data points to estimate the slope and intercept. These are then used to adjust the *iRec* estimate as shown on page 10. Simple linear-regression hypothesis tests that the slope is equal to 1 and the intercept equal to 0. A systematic approach is used to investigate the impact of creel-estimate quality on the regression estimates. Generally, the more data points used, the more "stable" the estimates appear.

Lastly, recommendations are made as to where to collect creel survey information to best calibrate the *iRec* survey.

It is unclear what will be done with negative estimates of the slope as this will lead to negative estimates of catch!

3.2 Use of ordinary least squares may be inappropriate.

The use of ordinary least-squares for both the regression fit, and hypothesis test may be inappropriate here because of the uncertainty in both the creel and *iRec* estimates. Indeed, I would like to see in Figures 2-6 the error bars in both directions for each data point. If the uncertainty in the *X*-direction is small, then ordinary least squares may potentially be ok, but I cannot judge this from the information presented. If the uncertainty in *X* (the creel estimates) is large, then you will be in an "error-in-variables" problem. Fortunately, if you an estimate of the standard error for each estimate in both directions, a simple Bayesian model can be fit without too much that will account for this (known) uncertainty in both directions.

By incorporating the uncertainty in both *X* and *Y*, the appropriate weighting is also automatically made for each pair. Points with low standard errors will automatically be given more weight than points with larger standard errors.

There is some issue on the consistency of the adjustment. At the model currently stands, the intercept for the fitted line could be different than zero. This could lead to an inconsistency in the adjusted values where the sum of adjusted values (e.g. from two different areas) is different than the adjustment applied to the sum from the two areas. To avoid this problem, the fitted line should be forced to go through the origin, i.e. . Any non-linear line (e.g. a curved line) would also have this inconsistency problem and so should be avoided. It is unclear if regression with unequal variation over the creel estimates is needed.

A fit on the log-scale for both variables should also be investigated because rather than constant difference between the two estimates, perhaps there is a constant multiplicative response bias. Of course, taking the log of 0 will be problematic – in those cases, these data points will have to be removed as not being informative. On the log scale, the calibration equation reduces to and estimates the log(multiplicative) factor between the two sources of data. By plotting data on the log-scale, any deviations from a constant multiplicative factor may also be apparent by curvature in the response.

The fitted lines include creel estimates with different quality levels (Table 1). Some care is need when including creel estimates that do not have complete coverage. By definition then, the estimates must be biased low and so the points corresponding to these surveys will be shifted to the left and so would tend to pull the regression line upwards. These estimates should be adjusted for undercoverage before being used, e.g. multiply by a simple correction factor for undercoverage prior to the fit.

Figure 5 and 6 also show a potential high leverage point that basically determines the best fit line. This of course is bad news for calibration as including/excluding a single data point dramatically changes the fit.

It would also be helpful in Figures 2 to 6 to show a confidence band around the slope for the uncertainty in the fitted response. This would provide a visual aid to indicate cases where there was evidence that the calibration line has a slope different than 1.

I'm also potentially concerned about the reuse of the same respondents for estimating catch for different species in the same area and across area. My feeling is that this problem may be inconsequential because of the high sample sizes and most respondents would not be fishing in

multiple areas, but this should be investigated. I would like to see some sample size numbers for individual *iRec* estimates in some of the plots.

Finally, the unit of analysis was the month-area-species combination from the creel and *iRec* survey. In some cases, this lead to many 0 estimates and small samples (poor creel quality). There is no reason why pooled area-month-species combination could be used to ensure that the point used for calibration is of suitable quality.

3.3 Yearly/species calibration estimates.

Given that the *iRec* survey is currently in the trial phase, separate calibration estimates for each year are likely appropriate. However, once the *iRec* survey becomes established, a single calibration for all year should be considered. Unfortunately, this would imply that past year estimates would have to be revised once a new survey-year is complete which is awkward. A solution for this would be use a weighted average of the calibration coefficients used last year and the new calibration estimates this year or to use a running average of the adjustment factors. This way the calibration numbers will have less variation over time, but still are capable to respond to changing biases over time.

The authors present separate slopes for each year, but do not do a formal analysis of changes over time. Without measures of precision for the fitted slopes, it is difficult to know if observed changes in slopes over time are not simply sampling artefacts. An analysis of covariance (adjusted for error-in-variables) could also be used to formally test if the slope is the same across years.

Similarly, different lines are fit for each species. Comparing the slopes among species would also be easier if precision estimates of the individual slopes were available. It is difficult to do a formal test of equality of slopes because the same area-month combinations are used, (e.g. for salmon species) and so a simple ANCOVA analysis will not be appropriate.

3.4 Designing creel survey to improve calibration.

The document concludes that the creel survey should continue to focus on area-month combinations with the highest effort and catch. As seen in Figures 2-6 this gives a wide-range in both the *X*- and *Y*- directions which is suitable for the regression approach. There does appears to be many month-area combinations with 0 catches (especially for the less popular species). As noted previously, these could be combined to get estimates at a larger scale (both temporally and spatially) to avoid so many 0 values during the calibration fit

3.5 Combining creel and *iRec* estimates.

The authors suggest finding yearly estimates using creel surveys when available plus *iRec* estimates for area-months when creel surveys are not run. Alternatively, the yearly estimates could be found using the adjusted *iRec* estimates for all area-month combinations. There is no statistical reason to choose one over the other.

Of course, the final standard error for the yearly estimates must be increased to account for the uncertainty in the adjustment process.

Editorial and other comments (Paper 1)

p.26 (and elsewhere). Distinguish between estimates and population values.

Equation (and others) fail to distinguish between population values and estimators of the same. For example, the first equation on page 26 is the estimator for the mean catch and the last equation on page 26 is an estimator of the variance of the ultimate units and the equation at the top of page 27 is the estimated variance of the mean catch per licence holder.

p.27. What df are use for confidence intervals?

What df are used in the t-multiplier for the confidence limits.

p. 31. "Understanding this composition is important because the analysis treats ...but reported a zero catch for the specified species across all days fished in that combination on the month". If I understand correctly, this is just a standard domain estimation method where a value of 0 is imputed for any respondent who does not have activity in the domain. Is there a desire for estimates per active respondent in the domain? If so, then this is just a ratio estimator of the response to an indicator variable 1/0 if the respondent was active in the domain. Cannot a question be asked to see if the respondent was active in the area at any time during the survey-month rather than relying on indirect measurement? This would at least distinguish respondents who were active and did not catch any fish, vs. respondents who were inactive and, by definition, captured no fish. A ratio-estimator is then easily computed to find the average catch per active angler.

p. 42. Automatic data flagging. I was rather surprised that no simple data editing is done when the data is captured, i.e. disallowing extremely large catches.

p. 43. The decision to be conservative in how to allocate sample rates to avoid "exhausting" licences can be overcome by doing the cluster survey as described above.

p. 43 "Our analysis approach is likely not providing appropriate precision statements for *iRec*". Some of these problems can be overcome by aggregating within respondents before finding estimates rather than aggregating estimates. The survey is design based, so concerns about normality are unnecessary – the estimated standard errors do not depend on the assumption of normality.

Multiple locations. Clarify "year". The term "year" is used interchangeable for fiscal-year and calendar year.

Editorial and other comments (Paper 2)

Figures 2 to 6. Use different plotting symbols for the different quality of creel interviews.

References.

Groves, R.M., Fowler Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). Hoboken, New Jersey: John Wiley and Sons.

Sarndal, C.-E. (2007). The calibration approach in survey theory and practice. Survey Methodology 33, 99-119.

Thompson, S. K. (2012). Sampling, 3rd Edition. Wiley: New York.

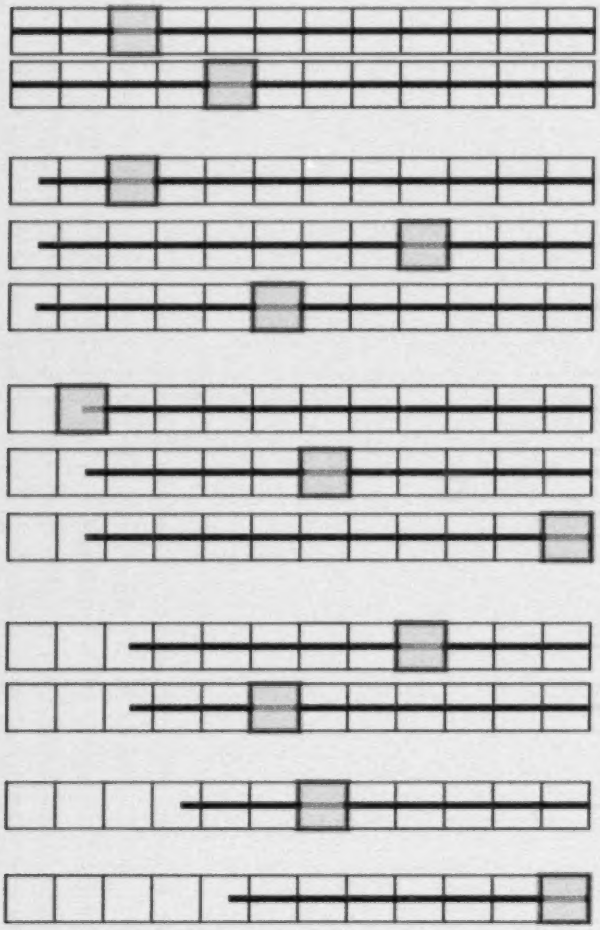

Figure 1. Conceptual diagram of two-stage cluster design. The licences (set of 12 boxes with black line indicating when the licence is active) are stratified by when purchased (and other attributes). From each licence a single month (box in grey) is selected with some probability design.